

A NEW APPROACH TO JORDAN DECOMPOSITION FOR FORMAL DIFFERENTIAL OPERATORS

MASOUD KAMGARPOUR AND SAMUEL WEATHERHOG

ABSTRACT. A theorem of Hukuhara, Levelt, and Turrittin states that every formal differential operator has a Jordan decomposition. We provide a new proof of this theorem by showing that every differential polynomial has a linear factorisation. The latter statement can be considered as a differential analogue of Puiseux's Theorem. Our approach makes clear the analogy between linear and differential operators thus making the proof more transparent.

CONTENTS

1. Introduction	1
2. Factorisation of differential polynomials	4
3. Proofs of the main results	8
References	11

1. INTRODUCTION

Let $\mathcal{K} := \mathbb{C}((t))$ be the field of formal Laurent series and consider the derivation $d : \mathcal{K} \rightarrow \mathcal{K}, d := t \frac{d}{dt}$. Let V be a finite dimensional vector space over \mathcal{K} . A formal differential operator is an additive map $D : V \rightarrow V$ satisfying the Leibniz rule

$$(1) \quad D(av) = aD(v) + d(a)v, \quad a \in \mathcal{K}, \quad v \in V.$$

Just as a linear operator encodes a system of linear equations, a differential operator encodes an ordinary differential equation. Thus, finding canonical forms for formal differential operators plays a crucial role in the theory of meromorphic differential equations; cf. [Var96].

A celebrated theorem of Hukuhara, Levelt, and Turrittin, states that every formal differential operator has a Jordan decomposition. This theorem has numerous applications in the theory of differential operators and other areas of mathematics, cf. [Kat70, Kat87, Luu15, BY15, KS16]. The existence of a Jordan decomposition for differential operators was first proved by Turrittin [Tur55], building on earlier work of Hukuhara [Huk41]. Turrittin's argument was rather complicated involving nine different cases. Subsequently, Levelt gave a more conceptual (albeit still not straightforward) proof and formulated the correct uniqueness statement [Lev75]. As a corollary, he concluded that every differential operator has an eigenvalue. Levelt asked for a direct proof of this corollary, noting that this would considerably simplify the proof of Jordan decomposition. Subsequently, several authors provided alternative approaches to this Theorem cf. [Was65, Mal79, Rob80, BV83, Pra83, vdPS03, Ked10]. However, as far as we know, Levelt's question has not been answered. The purpose of this note is to provide an elementary proof of the fact that every differential operator has an eigenvalue and use it to provide a simple proof of the existence of a Jordan decomposition, thus fulfilling Levelt's vision.

2010 *Mathematics Subject Classification.* 13N10.

Key words and phrases. Meromorphic ordinary differential equations, formal differential operators, Hukuhara-Levelt-Turrittin Theorem, Differential polynomials, Differential Hensel's Lemma, Newton polygons.

1.1. Matrix presentation and gauge transformation. In analogy with linear operators, differential operators have matrix presentations and it will be convenient to have these at our disposal. Let $D : V \rightarrow V$ be a differential operator. Choosing a basis for V/\mathcal{K} , we can represent D as an operator $d + A$ where A is an $n \times n$ matrix with values in \mathcal{K} . Changing the basis by an element $g \in GL_n(\mathcal{K})$ amounts to changing the operator $d + A$ to $d + g^{-1}Ag + g^{-1}dg$. Here dg denotes the matrix obtained by applying the derivation d to each entry of g . The map

$$A \mapsto g^{-1}Ag + g^{-1}dg$$

is called *gauge transformation*.

Remark 1. There is an alternative realisation of gauge transformation via conjugation in the extended loop group. Consider the loop algebra $\mathfrak{g}(\mathcal{K}) := \mathfrak{g} \otimes \mathcal{K}$ and the extended loop algebra $\hat{\mathfrak{g}} := \mathbb{C}d + \mathfrak{g}(\mathcal{K})$. Then one knows that this Lie algebra integrates to an algebraic group \hat{G} called the extended loop group [Kac90]. The conjugation action of \hat{G} on $\hat{\mathfrak{g}}$ restricts to give an action of $G(\mathcal{K})$ on $d + \mathfrak{g}(\mathcal{K})$ and therefore, also on $\mathfrak{g}(\mathcal{K})$. This action coincides with the gauge transformation defined above.

1.2. Main results. Before discussing Jordan decomposition, we have to explain what a semisimple differential operator is. A differential operator $D : V \rightarrow V$ is *simple* if it has no D -invariant subspace. It is *semisimple* if every D -invariant subspace has a D -invariant complement. It is easy to show that an operator is semisimple if and only if it is a direct sum of simple ones.

Semisimple differential operators have a straightforward description in terms of diagonal matrices. To see this, recall that a classical theorem of Puiseux states that for every positive integer b , $\mathcal{K}_b := \mathbb{C}((t^{\frac{1}{b}}))$ is the unique extension of \mathcal{K} of degree b . The derivation d extends canonically to a derivation d_b on \mathcal{K}_b . Moreover, given a differential operator D , one has a canonical differential operator

$$D \otimes_{\mathcal{K}} \mathcal{K}_b : V \otimes_{\mathcal{K}} \mathcal{K}_b \rightarrow V \otimes_{\mathcal{K}} \mathcal{K}_b$$

called the base change of D to \mathcal{K}_b . All base changes considered in this article are of this form.

Notation 2. Henceforth, we will use the notation $V_b := V \otimes_{\mathcal{K}} \mathcal{K}_b$ and $D_b = D \otimes_{\mathcal{K}} \mathcal{K}_b$.

Let us call a differential operator $D : V \rightarrow V$ *diagonalisable* if it has a presentation of the form $d + A$ where A is a diagonal matrix.

Theorem 3 (Levelt). *A formal differential operator is semisimple if and only if, after an appropriate finite base change, it is diagonalisable.*

Theorem 4 (Hukuhara-Levelt-Turrittin). *Every formal differential operator D can be written as a sum $D = S + N$ of a semisimple differential operator S together with a nilpotent \mathcal{K} -linear operator N such that S and N commute. Moreover, the pair (S, N) is unique.*

Using the classification of nilpotent linear operators, we can rephrase Theorem 4 to state that every differential operator has a *Jordan form*; cf. [Lev75, §6.(a)].

Remark 5. In the language of extended loop algebras (Remark 1), we can reformulate Theorem 4 as follows: after an appropriate base change, every element of $\hat{\mathfrak{g}}$ is \hat{G} -conjugate to one of the form $\lambda d + X + N$ where $\lambda \in \mathbb{C}$, $\lambda d + X$ is a semisimple λ -differential operator, and N is a nilpotent linear operator. From this point of view, it would be desirable to seek a classification of conjugacy classes of affine Kac-Moody algebras (including the twisted Kac-Moody algebras). For some results in this direction, cf. [BV83]. If instead of conjugacy classes in the extended loop algebra, we consider the conjugacy classes in the extended loop group, then we obtain the notion of *q-difference operator*. There is also a version of Jordan decomposition for these objects, cf. [Pra83].

1.3. Eigenvalues and eigenvectors. In analogy with linear operators, formal differential operators have eigenvalues and eigenvectors.

Definition 6. *An eigenvector for a formal differential operator $D : V \rightarrow V$ is a non-zero vector $v \in V$ satisfying $Dv = av$ for some $a \in \mathcal{K}$. In this case, a is called an eigenvalue.*

Let us make a digression to note an obvious difference between the linear and differential setting. If $Dv = av$ then

$$(2) \quad D(cv) = (a + c^{-1}d(c))(cv), \quad c \in \mathcal{K}.$$

This motivates the following definition.

Definition 7. We call $a, b \in \mathcal{K}$ similar if $a - b = c^{-1}d(c)$ for some $c \in \mathcal{K}$.

It is easy to show that this defines an equivalence relation on \mathcal{K} . It follows from (2) that if $a \in \mathcal{K}$ is an eigenvalue of D , then every element similar to a is also an eigenvalue. Thus, since \mathcal{K} is infinite, a differential operator may have infinitely many eigenvalues. The saving grace is that the set of eigenvectors for each eigenvalue forms a vector space over the subfield $\mathcal{K}^d \subseteq \mathcal{K}$ where $\mathcal{K}^d := \{x \in \mathcal{K} \mid d(x) = 0\}$. This is in contrast with the linear setting where an eigenspace is a vector space *over* \mathcal{K} .

Returning to our main point, we observe that Theorems 3 and 4 imply immediately the corollary alluded to in the opening paragraphs:

Corollary 8. Every formal differential operator has, after an appropriate finite base change, an eigenvalue.

As previously mentioned, we shall first prove this corollary and then deduce Theorems 3 and 4. We now discuss our approach in more detail.

1.4. Outline of our approach. Let $\mathcal{K}\{x\}$ denote the non-commutative ring of *differential polynomials*. As an abelian group $\mathcal{K}\{x\} = \mathcal{K}[x]$ but multiplication is modified by the rule $xa = ax + da$ for all $a \in \mathcal{K}$. In §2, we prove a version of Hensel's lemma for differential polynomials. We then use this to prove that every differential polynomial with coefficients in $\mathcal{O} := \mathbb{C}[[t]]$ has a linear factorisation over \mathcal{O}^1 . By making use of the Newton polygon of a differential polynomial, we conclude a key result in our approach:

Theorem 9. Every differential polynomial in $\mathcal{K}\{x\}$ has a linear factorisation over some finite extension.

We note that the above theorem appears in the literature as a corollary of Theorem 4. In our approach, however, we prove it directly.

In §3, we use Theorem 9 to prove our main results. As we shall see, Corollary 8 and Theorem 3 follow more or less immediately. Using these results, we obtain a generalised eigenspace decomposition for differential operators. This shows that every differential operator has a representation $d + X$ where X is a block-upper triangular matrix and each block has a unique (up to similarity) eigenvalue.

At this point, we encounter a subtle difference between the linear and differential setting. Let us write $X = Y + Z$ where Y is diagonal and Z is strictly upper triangular. If we were considering linear operators, then Y would be the semisimple operator and Z would be the nilpotent part of X and these two commute. In the differential setting, however, the situation is more subtle because the operators $d + Y$ and Z do not necessarily commute. In fact, these two operators commute if and only if the entries of Z are complex numbers (i.e. have no powers of t). We prove that indeed we can arrange so that elements of Z are complex numbers.

1.5. (Counter-)Examples. It is immediate from the definition that if $L : V \rightarrow V$ is a linear operator, then $D := d + L$ is a differential operator. Note, however, that the relationship between the Jordan decompositions of L and D is not straightforward. Indeed, this is one reason Theorems 3 and 4 are non-trivial for otherwise one could try deducing them from the corresponding results in the linear setting. We now give some examples illustrating this difference between the linear and differential setting.

Example 10. For each integer k , consider the linear operator

$$L_k := \begin{pmatrix} 0 & t^k \\ 0 & 0 \end{pmatrix} : \mathcal{K}^2 \rightarrow \mathcal{K}^2.$$

If $k \neq 0$, then

$$(d + L_k) \cdot \begin{pmatrix} -t^k \\ k \end{pmatrix} = (d + L_k) \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0.$$

Thus, we have two independent eigenvectors with eigenvalue 0. We conclude that when $k \neq 0$, the differential operator $d + L_k$ is semisimple while the linear operator L_k is nilpotent.

¹Note that this is stronger than what we have in the linear setting; see Remark 14.

Example 11. Conversely, we have semisimple linear operators whose associated differential operator is not semisimple. For instance, consider the linear operator

$$S_k := \begin{pmatrix} 1 & 0 \\ 1 & 1+t^k \end{pmatrix} : \mathcal{K}^2 \rightarrow \mathcal{K}^2.$$

This linear operator has distinct eigenvalues and is therefore semisimple. If $k \geq 0$, then one can show that $d + S_k$ is gauge equivalent to $d + \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$; thus, it is not semisimple (since its Jordan decomposition has a nontrivial nilpotent part).

The only relationship we know of between Jordan decompositions for a linear operator and its associated differential operator is for diagonal ones: it follows from Theorem 3 that if L is a linear operator on \mathcal{K}^n specified by a diagonal matrix, then $d + L$ is a semisimple differential operator.

1.6. Acknowledgment. We thank Philip Boalch, Peter McNamara, Daniel Sage, Ole Warnaar, and Sinead Wilson for helpful conversations. We are grateful to Claude Sabbah for sending copies of [Mal79] and [Sab]. The material in this paper forms a part of SW's Master's Thesis. MK was supported by an ARC DECRA Fellowship.

2. FACTORISATION OF DIFFERENTIAL POLYNOMIALS

The goal of this section is to prove Theorem 9. We start by introducing some notation. Let R be a \mathbb{C} -algebra and $d : R \rightarrow R$ a derivation. We let $R\{x, d\}$ denote the ring of differential polynomials over (R, d) . We will be interested in the case that $R = \mathcal{O} := \mathbb{C}[[t]]$ or $R = \mathcal{K} := \mathbb{C}((t))$ and the derivation d is of the form $\delta_m := t^m \frac{d}{dt}$, for some positive integer m . According to [Ore33], the ring $\mathcal{K}\{x, \delta_m\}$ is a left and right principal ideal domain.

2.1. Differential Hensel's Lemma. Let $f \in \mathcal{O}\{x, \delta_m\}$ be a differential polynomial. We write $f \pmod{t^n}$ for the polynomial obtained by first moving all factors of t to the left and then reducing the coefficients modulo t^n . We denote $f \pmod{t}$ by \bar{f} . Note that this is a polynomial in $\mathbb{C}[x]$. Without the loss of generality, we assume throughout that $\bar{f} \neq 0$.

Now suppose we have a factorisation of the form

$$\bar{f} = g_0 h_0, \quad g_0, h_0 \in \mathbb{C}[x].$$

Our aim is to lift this to a factorisation of f in $\mathcal{K}\{x, \delta_m\}$. We think of the following result as a differential analogue of Hensel's lemma.

Proposition 12. *Let $f \in \mathcal{O}\{x, \delta_m\}$ and $\bar{f} = g_0 h_0$ as above. Suppose*

$$\begin{cases} \gcd(g_0(x+n), h_0(x)) = 1, & \forall n \in \mathbb{Z}_{>0} & \text{if } m = 1 \\ \gcd(g_0(x), h_0(x)) = 1 & & \text{if } m > 1. \end{cases}$$

Then we have a factorisation $f = gh$ with $g, h \in \mathcal{O}\{x, \delta_m\}$, $\deg(g) = \deg(g_0)$, $\bar{g} = g_0$ and $\bar{h} = h_0$.

We note that a version of this proposition appeared in [Pra83, Lemma 1].

Proof. First of all, in the differential polynomial ring $\mathcal{K}\{x, \delta_m\}$, easy induction arguments show that

$$(3) \quad h(x)t^i = t^i h(x + it^{m-1}), \quad \forall h(x) \in \mathcal{K}\{x, \delta_m\}, \quad \forall i \in \mathbb{Z},$$

and

$$(4) \quad (t^d x)^k = \sum_{j=0}^{k-1} a_j t^{kd+(m-1)j} x^{k-j}, \quad \forall d \in \mathbb{Z} - \{0\}, \quad \forall k \in \mathbb{N},$$

for some constants $a_j \in \mathbb{C}$, $a_0 = 1$.

Our goal is to inductively build a sequence of polynomials

$$(5) \quad g_n(x) = g_0 + t p_1 + t^2 p_2 + \cdots + t^{n-1} p_{n-1} + t^n p_n, \quad p_i \in \mathbb{C}[x]$$

$$(6) \quad h_n(x) = h_0 + t q_1 + t^2 q_2 + \cdots + t^{n-1} q_{n-1} + t^n q_n, \quad q_i \in \mathbb{C}[x],$$

which satisfy:

$$f \equiv g_n(x)h_n(x) \pmod{t^{n+1}}.$$

If we can do this, then by letting $n \rightarrow \infty$ we will obtain elements $g, h \in \mathcal{O}\{x, \delta_m\}$ such that $f = gh$.

Suppose that we know the p_i and q_i for $1 \leq i \leq n-1$. In view of (5) and (6) we have:

$$g_n = g_{n-1} + t^n p_n, \quad h_n = h_{n-1} + t^n q_n.$$

Requiring that $f \equiv g_n(x)h_n(x) \pmod{t^{n+1}}$ then gives us the following condition:

$$\begin{aligned} f &\equiv g_n(x)h_n(x) && \pmod{t^{n+1}} \\ &\equiv (g_{n-1}(x) + t^n p_n(x))(h_{n-1}(x) + t^n q_n(x)) && \pmod{t^{n+1}} \\ &\equiv g_{n-1}(x)h_{n-1}(x) + g_{n-1}(x)t^n q_n(x) + t^n p_n(x)h_{n-1}(x) + t^n p_n(x)t^n q_n(x) && \pmod{t^{n+1}}. \end{aligned}$$

We need to shift the powers of t to the left. By (3), $g_{n-1}(x)t^n = t^n g_{n-1}(x + nt^{m-1})$, so we have:

$$\begin{aligned} f - g_{n-1}(x)h_{n-1}(x) &\equiv t^n g_{n-1}(x + nt^{m-1})q_n(x) + t^n p_n(x)h_{n-1}(x) && \pmod{t^{n+1}} \\ &\equiv t^n (g_{n-1}(x + nt^{m-1})q_n(x) + p_n(x)h_{n-1}(x)) && \pmod{t^{n+1}}, \end{aligned}$$

and thus

$$\begin{aligned} \frac{f - g_{n-1}(x)h_{n-1}(x)}{t^n} &\equiv g_{n-1}(x + nt^{m-1})q_n(x) + p_n(x)h_{n-1}(x) && \pmod{t} \\ &\equiv g_0(x + nt^{m-1})q_n(x) + p_n(x)h_0(x) && \pmod{t}. \end{aligned}$$

For notational convenience, we set:

$$f_n = \frac{f - g_{n-1}(x)h_{n-1}(x)}{t^n}.$$

so that we have

$$(7) \quad f_n \equiv g_0(x + nt^{m-1})q_n(x) + p_n(x)h_0(x) \pmod{t}.$$

Now if $m > 1$, then (7) reduces to

$$f_n \equiv g_0(x)q_n(x) + p_n(x)h_0(x) \pmod{t}.$$

Since $\mathbb{C}[x]$ is a Euclidean domain, we will be able to solve this for p_n and q_n provided that g_0 and h_0 are coprime. On the other hand, if $m = 1$, then (7) becomes

$$f_n \equiv g_0(x + n)q_n(x) + p_n(x)h_0(x) \pmod{t}.$$

In this case, we will only be able to generate the entire sequence if $g_0(x + n)$ and $h_0(x)$ are coprime for all $n \in \mathbb{Z}_{>0}$.

All that remains to show is that we can control the degree of the g_n 's. We will show this in the case $m = 1$. The proof in the case $m > 1$ is similar (replace $g_0(x + n)$ with $g_0(x)$ everywhere). Since $g_0(x + n)$ and $h_0(x)$ are coprime, we can find $a, b \in \mathbb{C}[x]$ such that

$$g_0(x + n)a(x) + h_0(x)b(x) = 1.$$

Multiplying through by f_n yields

$$(8) \quad g_0(x + n)a(x)f_n(x) + h_0(x)b(x)f_n(x) = f_n(x).$$

Using the division algorithm we can find unique p_n and q_n such that $\deg(p_n) < \deg(g_0)$. Write:

$$b(x)f_n(x) = Q(x)g_0(x) + R(x)$$

with $\deg(R) < \deg(g_0)$. Equation (8) then becomes:

$$g_0(x + n)(a(x)f_n(x) + Q(x)h_0(x)) + h_0(x)R(x) \equiv f_n(x) \pmod{t}.$$

Setting $p_n = R$ and $q_n = af_n + Qh_0$ gives us the required g_n and h_n . □

Corollary 13. *Let $f \in \mathcal{O}\{x, \delta_1\}$ be a monic differential polynomial. Then f admits a factorisation of the form*

$$(x - \Lambda)h,$$

with $\Lambda \in \mathcal{O}$ and $h \in \mathcal{O}\{x, \delta_1\}$.

Proof. Let $\bar{f} \in \mathbb{C}[x]$ be the reduction of $f \bmod t$. Since f is monic, \bar{f} is non-constant and hence factors over \mathbb{C} into linear factors:

$$\bar{f} = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_n), \quad \lambda_i \in \mathbb{C}.$$

Without loss of generality, we can order these factors so that $\operatorname{Re}(\lambda_1) \leq \operatorname{Re}(\lambda_2) \leq \cdots \leq \operatorname{Re}(\lambda_n)$. With this ordering we then have

$$\bar{f} = g_0 h_0,$$

where

$$g_0 = x - \lambda_1, \quad h_0 = (x - \lambda_2) \cdots (x - \lambda_n).$$

By our choice of ordering, $g_0(x + n)$ has no common factor with h_0 for all $n \in \mathbb{Z}_{>0}$. Hence we can apply Proposition 12 to obtain a factorisation of the form

$$f = (x - \Lambda)h, \quad \Lambda \in \mathcal{O}, h \in \mathcal{O}\{x, \delta_1\},$$

as required. \square

Remark 14. Note that the above result is false for the usual polynomial ring $\mathcal{O}[x]$. Indeed, $x^2 + t - t^2$ does not have a linear factorisation over this ring, but if we consider it as an element of $\mathcal{O}\{x, \delta_1\}$, then $x^2 + t - t^2 = (x - t)(x + t)$.

2.2. From power series to Laurent series via Newton polygons. In the previous section, we settled linear factorisation for differential polynomials in $\mathcal{O}\{x, \delta_1\}$. In this section, we explain how, by a change of variable, we can transform polynomials with coefficients in \mathcal{K} to those with power series coefficients. The price is that we have to go to a finite extension \mathcal{K}_q of \mathcal{K} and, more seriously, the derivation is not simply the canonical extension of δ_1 to \mathcal{K}_q . Nevertheless, we shall see that this change of variable allows us to factor elements of $\mathcal{K}\{x, \delta_1\}$.

Lemma 15. *Consider the differential polynomial $f(x) = \sum_{i=0}^n a_i x^{n-i} \in \mathcal{K}\{x, \delta_1\}$. Let $r := \min \{ \frac{v(a_i)}{i} \}$. Then $g(X) = t^{-nr} f(t^r X)$ is a monic differential polynomial with power series coefficients.*

Proof. To be more precise, write $r = \frac{p}{q}$ with $\gcd(p, q) = 1$ and $q > 0$. Since we have already dealt with the case $f \in \mathcal{O}\{x, \delta_1\}$, we may assume that $r < 0$. In order to make the change of variables $x = t^r X$, we require a field extension to $\mathcal{K}_q = \mathbb{C}((t^{1/q}))$. Let $s := t^{1/q}$ so that our change of variables becomes $x = s^p X$. Applying (4) to $f(s^p X)$ yields $f(s^p X) = s^{np} g(X)$ where

$$g(X) = s^{-np} a_n + \sum_{k=0}^{n-1} a_k \sum_{j=0}^{n-1-k} m_{n-k,j} s^{(-j-k)p} X^{n-1-k-j}.$$

Since $v_s(a_i) = qv_t(a_i)$, $v_t(a_i) \geq \frac{ip}{q}$ implies that $v_s(a_i) \geq ip$. Thus, for $0 \leq l \leq n-1$, the coefficient, b_l , of X^{n-l} in g satisfies

$$v_s(b_l) = \min_{0 \leq k \leq l} \{v_s(a_k s^{-lp})\} \geq \min_{0 \leq k \leq l} \{kp - lp\} = 0,$$

where the last equality follows since $p < 0$.

It is clear that $v_s(b_l)$ will be 0 exactly when $v_s(a_l) = lp$, that is, if, and only if, $v_t(a_l) = lr$. For the “constant” term of g we have

$$v_s(b_n) = v_s(a_n s^{-np}) \geq np - np = 0,$$

again with equality exactly when $v_t(a_n) = nr$. Thus

$$g(X) = X^n + b_1 X^{n-1} + \cdots + b_n, \quad b_i \in \mathbb{C}[[s]],$$

with $\min(v_s(b_i)) = 0$. Furthermore, $v_s(b_i) = 0$ if, and only if, $v_t(a_i) = ir$. This shows that $g(X) \in \mathbb{C}[[s]]\{X, \frac{1}{q}s^{1-p}\frac{d}{ds}\}$ (note the change of derivation). \square

Consider $g(x)$ from the above lemma. If $\bar{g}(x)$ has two distinct roots, then Hensel’s lemma allows us to factor it. We now study the opposite extreme, i.e., when all roots of $\bar{g}(x)$ are equal. It will be helpful to use the notion of Newton polygon for differential polynomials, cf. [Ked10, §6.4].

Definition 16 (Newton Polygon). Let $f \in \mathcal{K}\{x, \delta_m\}$ be a differential polynomial and write

$$f(x) = \sum_{i=0}^n a_i x^{n-i}, \quad a_i \in \mathcal{O}.$$

Consider the lower boundary of the convex hull of the points

$$\{(n-i), v_t(a_i) : 0 \leq i \leq n\} \subset \mathbb{R}^2.$$

The Newton polygon of f , denoted $NP(f)$, is obtained from this boundary by replacing all line segments of slope less than $1-m$ with a single line segment of slope exactly $1-m$.

Lemma 15 now has the following corollary.

Corollary 17. Let f and g be as in Lemma 15 and suppose that $\bar{g} := g \pmod{s} = (X + \lambda)^n$, $\lambda \in \mathbb{C}$. Then λ is non-zero and the Newton polygon of f has a single integral slope.

Proof. As in Lemma 15, write

$$g(X) = X^n + b_1 X^{n-1} + \cdots + b_n, \quad b_i \in \mathbb{C}[[s]].$$

Since $\min\{v_s(b_i)\} = 0$, $\lambda \neq 0$. Now since, $\lambda \neq 0$, expanding the bracket $(X + \lambda)^n$ shows that $v_s(b_i) = 0$ for all i and hence $v_t(a_i) = ir$. Thus, the Newton polygon of f has a single slope of r and since $v_t(a_1) = r$, r is an integer. \square

For future use, we also record the following lemma.

Lemma 18. Let f and g be as in Lemma 15 and suppose that $\bar{g} = (X + \lambda)^n$, $\lambda \in \mathbb{C}$. Then the Newton polygon of $f(x - \lambda t^r)$ has a single slope strictly smaller than the slope of the Newton polygon of $f(x)$.

Proof. By Corollary 17, r is an integer and hence no extension of \mathcal{K} is necessary. Since $\bar{g} = (X + \lambda)^n$, we can write g as

$$g = (X + \lambda)^n + e_1(X + \lambda)^{n-1} + \cdots + e_n, \quad e_i \in \mathcal{O},$$

with $v_t(e_i) > 0$ for all i . Now

$$\begin{aligned} f(t^r X) &= t^{nr} ((X + \lambda)^n + e_1(X + \lambda)^{n-1} + \cdots + e_n) \\ \implies f(x) &= t^{nr} ((t^{-r}x + \lambda)^n + e_1(t^{-r}x + \lambda)^{n-1} + \cdots + e_n), \end{aligned}$$

and hence

$$f(x - \lambda t^r) = t^{nr} ((t^{-r}x)^n + e_1(t^{-r}x)^{n-1} + \cdots + e_n).$$

Applying (4), we have, for $m_{k,l} \in \mathbb{C}$,

$$\begin{aligned} f(x - \lambda t^r) &= t^{nr} \left(t^{-nr} \sum_{j=0}^{n-1} m_{n,j} x^{n-j} + e_1 t^{-(n-1)r} \sum_{j=0}^{n-2} m_{n-1,j} x^{n-1-j} + \cdots + e_n \right) \\ &= \sum_{j=0}^{n-1} m_{n,j} x^{n-j} + e_1 t^r \sum_{j=0}^{n-2} m_{n-1,j} x^{n-1-j} + \cdots + t^{nr} e_n. \end{aligned}$$

Since $v(e_i) > 0$, the valuation of the coefficient of x^{n-j} in $f(x - \lambda t^r)$ is strictly greater than the corresponding coefficient in $f(x)$. This means that the slope of the Newton polygon for $f(x - \lambda t^r)$ is strictly less than the slope of the Newton polygon for $f(x)$. \square

Example 19. In order to illustrate Corollary 17 and Lemma 18, consider the differential polynomial

$$f_1(x) = x^2 + (4t^{-2} + 2t^{-1} + 2)x + (4t^{-4} + 4t^{-3} + t^{-2} + t^{-1} + 1).$$

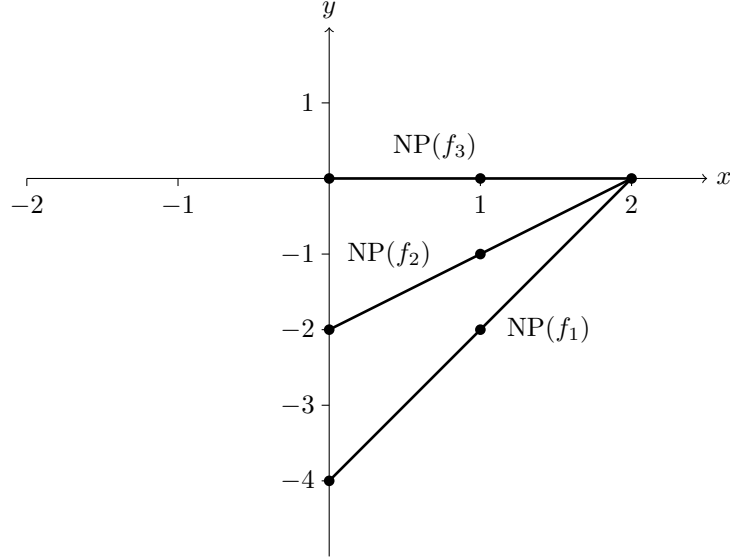
In this case, $r = -2$ and the change of variables $x = t^{-2}X$ yields

$$g_1(X) = X^2 + (4 + 2t)X + (4 + 4t + t^2 + t^3 + t^4),$$

and so $\bar{g}_1(X) = (X + 2)^2$. The figure below shows that the Newton polygon of f_1 has only a single slope of -2 (cf. Cor 17). Making the translation $x \mapsto x + 2t^{-2}$ as in Lemma 18 yields the new polynomial

$$f_2(x) = x^2 + (2t^{-1} + 2)x + (t^{-2} + t^{-1} + 1).$$

This has a single slope $r = -1$ and a final translation $x \mapsto x - t^{-1}$ yields $f_3(x) = x^2 + 2x + 1$. This can easily be factorised and reversing the change of variables yields the full factorisation $f_1(x) = (x + 2t^{-2} + t^{-1} + 1)^2$.



2.3. Proof of Theorem 9. Write $f(x) = x^n + a_1x^{n-1} + \dots + a_n \in \mathcal{K}\{x, \delta_1\}$ and let $r := \min \left\{ \frac{v(a_i)}{i} \right\} \in \mathbb{Q}$. If $r \geq 0$, then the result follows from the differential Hensel's Lemma (see Corollary 13) so we may assume $r < 0$. Let us write

$$r = \frac{p}{q}, \quad q > 0, \quad \gcd(p, q) = 1.$$

Consider the transformation $x \mapsto t^r X$. Note that the differential field (\mathcal{K}, δ_1) extends to $(\mathcal{K}_q, \frac{1}{q}s^{1-p}\frac{d}{ds})$ where $s := t^{1/q}$. Moreover, we obtain a monic differential polynomial $g(X) \in \mathbb{C}[[s]]\{y, \frac{1}{q}s^{1-p}\frac{d}{ds}\}$. Let $\bar{g}(X)$ denote the reduction of $g(X)$ modulo the maximal ideal of $\mathbb{C}[[s]]$. If $\bar{g}(X)$ has two distinct roots, then we can again apply Proposition 12 to reduce the problem to a polynomial of degree strictly less than f . Thus, we are reduced to the case that $\bar{g}(X)$ has a unique repeated root λ . For inductive purposes, we rename f to f_1 . In this case, by Corollary 17, $\lambda \neq 0$ and the Newton polygon of f_1 has a single *integral* slope. Now we make the transformation $x \mapsto x - \lambda t^r$. As shown in Lemma 18, under this transformation f_1 is mapped to a polynomial f_2 whose Newton polygon has a single slope strictly less than that of f_1 . Note that this transformation does not change the differential field.

Now we start the process with the polynomial $f_2(x) := x^n + b_1x^{n-1} + \dots + b_n \in \mathcal{K}\{x, \delta_1\}$; i.e., we let $r_2 := \min \left\{ \frac{v(b_i)}{i} \right\}$. If $r_2 \geq 0$ we are done. Otherwise, we make the change of variable $x \mapsto t^{r_2}X$ to obtain a new polynomial $g_2(X)$. If $\bar{g}_2(X)$ has distinct roots, then we are done; otherwise, applying Corollary 17 again, we conclude that the Newton polygon of f_2 has a single *integral* slope. Since the slope of f_2 is a nonnegative integer strictly less than slope of f_1 , this process must stop in finitely many steps at which point we have a factorisation of our polynomial. □

3. PROOFS OF THE MAIN RESULTS

3.1. Proof of Corollary 8 (Every differential operator has an eigenvalue). The argument proceeds exactly as in the linear setting. Let $D : V \rightarrow V$ be a differential operator and $v \in V$ be a non-zero vector. Consider the sequence $v, D(v), D^2(v), \dots$. As V has finite dimension over \mathcal{K} , we must have that

$$D^n(v) + a_1D^{n-1}(v) + \dots + a_{n-1}D(v) + a_nv = 0, \quad a_i \in \mathcal{K},$$

where $n = \dim_{\mathcal{K}}(V)$. Now consider the polynomial $f(x) = x^n + a_1x^{n-1} + \dots + a_n$ in the twisted polynomial ring $\mathcal{K}\{x\}$. After a finite extension, we can write

$$f(x) = (x - \Lambda_1) \cdots (x - \Lambda_n) \in \mathcal{K}_b\{x\}, \quad \Lambda_i \in \mathcal{K}_b, b \in \mathbb{Z}_{>0}.$$

Thus,

$$(D_b - \Lambda_1) \cdots (D_b - \Lambda_n)v = 0.$$

Let $i \in \{1, 2, \dots, n\}$ be the largest number such that $(D_b - \Lambda_i) \cdots (D_b - \Lambda_n)v = 0$. If $i = n$, then v is an eigenvector of D_b with eigenvalue Λ_n . Otherwise $(D_b - \Lambda_{i+1}) \cdots (D_b - \Lambda_n)v$ is an eigenvector of D_b with eigenvalue Λ_i . \square

3.2. Proof of Theorem 3 (Semisimple operators are diagonalisable). We need the following lemma. The proof is an easy argument using the Galois group $\text{Gal}(\mathcal{K}_b/\mathcal{K})$; see [Lev75, §1(e)] for details.

Lemma 20. *D is semisimple if and only if D_b is.*

Now we are ready to prove Theorem 3. Suppose D is semisimple. We prove by induction on $\dim(V)$ that, after an appropriate base change, it is diagonalisable. If $\dim(V) = 1$ the result is obvious, so assume $\dim(V) > 1$. Without the loss of generality, assume D has an eigenvalue v (if not, do an appropriate base change; by the previous lemma, the operator remains semisimple). Let $U = \text{span}_{\mathcal{K}}\{v\}$. Then U is a one-dimensional, D -invariant subspace of V ; thus, there exists a D -invariant complement W . Now $D : W \rightarrow W$ is semisimple so by our induction hypothesis (after an appropriate base change), we can write W as a direct sum of one-dimensional subspaces. Thus, after an appropriate base change, we have a decomposition of our vector space into one-dimensional, invariant subspaces and so D is diagonalisable.

Conversely, suppose D_b is a diagonalisable operator. Then clearly D_b is semisimple. By the previous lemma, so is D . \square

3.3. Generalised eigenspace decomposition. Let $D : V \rightarrow V$ be a formal differential operator and let $a \in \mathcal{K}$.

Definition 21. *The generalised eigenspace $V(a)$ of D is defined as*

$$V(a) := \text{span}_{\mathcal{K}}\{v \in V \mid (D - a)^n = 0, \text{ for some positive integer } n.\}$$

The goal of this section is to prove the following theorem.

Theorem 22 (Generalised eigenspace decomposition). *There exists a finite extension $\mathcal{K}_b/\mathcal{K}$ such that we have a canonical decomposition $V_b = \bigoplus_i V_b(a_i)$. Moreover,*

$$V_b(a_i) \cap V_b(a_j) \neq \{0\} \iff a_i \text{ is similar to } a_j \iff V_b(a_i) = V_b(a_j).$$

Before proving this theorem, we need to recall some facts about differential operators. Let $D : V \rightarrow V$ be a differential operator. Define

$$\begin{aligned} H^0(V) &:= \ker(D), \\ H^1(V) &:= V/D(V). \end{aligned}$$

Note that these are vector spaces over \mathbb{C} (not over \mathcal{K}). The following proposition due to Malgrange [Mal74, Theorem 3.3] is an analogue of the rank-nullity theorem for formal differential operators.

Proposition 23. *Let $D : V \rightarrow V$ be a formal differential operator. Then*

$$\dim_{\mathbb{C}} H^0(V) = \dim_{\mathbb{C}} H^1(V).$$

Next, recall that the dual differential operator $D : V \rightarrow V$ is the operator D^* on the vector space $V^* = \text{Hom}_{\mathcal{K}}(V, \mathcal{K})$ defined by

$$D^* : V^* \rightarrow V^*, \quad D^*(f) = f \circ D - d \circ f, \quad f \in V^*.$$

Let $D : V \rightarrow V$ and $D' : V' \rightarrow V'$ be differential operators. Then, we can define a differential operator $D \otimes D'$ on $V \otimes V'$ by

$$(D \otimes D')(v \otimes v') := D(v) \otimes v' + v \otimes D'(v').$$

The set of all of $\mathcal{K}\{x\}$ -linear maps from V to V' is denoted $\text{Hom}_{\mathcal{K}\{x\}}(V, V')$. This is a \mathbb{C} -vector space. The Yoneda extension group $\text{Ext}_{\mathcal{K}\{x\}}^1(V, V')$ consists of equivalence classes of extensions of $\mathcal{K}\{x\}$ -modules

$$0 \rightarrow V \rightarrow V'' \rightarrow V' \rightarrow 0$$

As usual, two extensions are equivalent if there exists a $\mathcal{K}\{x\}$ -linear isomorphism between them inducing the identity on V and V' .

Proposition 24. *Let $D : V \rightarrow V$ and $D' : V' \rightarrow V'$ be two formal differential operators. Then, we have*

- (i) $\dim_{\mathbb{C}} \operatorname{Ext}_{\mathcal{K}\{x\}}^1(V, V') = \dim_{\mathbb{C}} H^0(V^* \otimes V')$.
- (ii) *If no eigenvalue of D is similar to an eigenvalue of D' , then $\operatorname{Ext}_{\mathcal{K}\{x\}}^1(V, V') = 0$.*

Proof. One can show (see [Ked10, Lemma 5.3.3]) that there is a canonical isomorphism of \mathbb{C} -vector spaces:

$$\operatorname{Ext}_{\mathcal{K}\{x\}}^1(V, V') \simeq H^1(V^* \otimes V').$$

This fact together with Proposition 23 implies (i).

The eigenvalues of $D^* \otimes D'$ are of the form $-a + a'$ where a and a' are eigenvalues of D and D' , respectively. By assumption, $-a + a'$ is never similar to zero; thus, kernel of $D^* \otimes D'$ is trivial. Part (ii) now follows from Part (i). \square

Proof of Theorem 22. We may assume, without the loss of generality, that all eigenvalues of D are already in \mathcal{K} (if not, do an appropriate base change). We use induction on $\dim(V)$ to prove the theorem. If $\dim(V) = 1$ then the claim is trivial. Suppose $\dim(V) > 1$. Then by assumption D has an eigenvector. Hence, we have a one-dimensional invariant subspace $U \subset V$. Let $W := V/U$. Then D defines a differential operator on W . Moreover, $V \in \operatorname{Ext}_{\mathcal{K}\{x\}}^1(U, W)$. By induction we may assume that W decomposes as

$$W = \bigoplus_i W(a_i), \quad a_i \in \mathcal{K},$$

for non-similar a_i . Now

$$V \in \operatorname{Ext}_{\mathcal{K}\{x\}}^1\left(U, \bigoplus_i W(a_i)\right) \simeq \bigoplus_i \operatorname{Ext}_{\mathcal{K}\{x\}}^1(U, W(a_i)).$$

If the eigenvalue a of $D|_U$ is not similar to any a_i then by the above proposition all the extension groups are zero, and so $V = W \oplus U$ and the theorem is established. If a is similar to a_j , for some j , then the only non-trivial component in the above direct sum is $\operatorname{Ext}_{\mathcal{K}\{x\}}^1(U, W(a_j))$. But it is easy to see that all differential operators in $\operatorname{Ext}_{\mathcal{K}\{x\}}^1(U, W(a_j))$ have only a single eigenvalue a_j (up to similarity). Hence V has the required decomposition. \square

3.4. Unipotent differential operators. Theorem 22 implies that we only need to prove Jordan decomposition for differential operators with a unique eigenvalue. By translating if necessary, we can assume this eigenvalue is zero. Thus, we arrive at the following:

Definition 25. *A differential operator is unipotent if all of its eigenvalues are similar to zero.*

We now give a complete description of unipotent differential operators. Let $\operatorname{Nilp}_{\mathbb{C}}$ denote the category whose objects are pairs (V, N) where V is a \mathbb{C} -vector space and N is a nilpotent endomorphism. The morphisms of $\operatorname{Nilp}_{\mathbb{C}}$ are linear maps which commute with N . Let \mathcal{U} be the category of pairs (V, D) consisting of a vector space V/\mathcal{K} and a unipotent differential operator $D : V \rightarrow V$. Define a functor

$$F : \operatorname{Nilp}_{\mathbb{C}} \rightarrow \mathcal{U}, \quad (V, N) \mapsto (\mathcal{K} \otimes_{\mathbb{C}} V, d + N).$$

The following result appears (without proof) in [Kat87, §2].

Lemma 26. *The functor F defines an equivalence of categories with inverse given by*

$$G : \mathcal{U} \rightarrow \operatorname{Nilp}_{\mathbb{C}}, \quad (V, D) \mapsto (\ker(D^{\dim_{\mathcal{K}}(V)}), D).$$

Proof. We first show that the composition $G \circ F$ equals the identity. Let $(V, N) \in \operatorname{Nilp}_{\mathbb{C}}$ with $n := \dim(V)$ and consider $F(V, N) = (V \otimes \mathcal{K}, d + N)$. The kernel of the operator $(d + N)^n$ acting on $V \otimes \mathcal{K}$ is the set of all constant vectors. This is an n -dimensional \mathbb{C} -vector space. Since d acts as 0 on this space, applying G to $(\mathcal{K} \otimes V, d + N)$ recovers the pair (V, N) .

Next, let $D : V \rightarrow V$ be a unipotent differential operator and let $n := \dim_{\mathcal{K}}(V)$. We first show by induction that $\ker(D^n)$ contains n \mathcal{K} -linearly independent vectors. If $n = 1$ this is obvious. If $n > 1$, then there exists $v \in V$ such that $Dv = 0$. Set $U := \operatorname{span}_{\mathcal{K}}\{v\}$ and consider the differential module V/U . This

has dimension $n - 1$ so we may assume there exist $\{v_1, \dots, v_n\}$ \mathcal{K} -linearly independent vectors in $\ker(D^{n-1})$. For each v_i we have $D^{n-1}v_i + U = U$ and hence $D^{n-1}v_i = a_i v$ for some $a_i \in \mathcal{K}$. Now observe that we can choose b_i such that $d^{n-1}(b_i) = a_i - a_{i,0}$ where $a_{i,0}$ is the constant term of a_i ; since we can always “integrate” elements with no constant term. Now we have

$$D^{n-1}(v_i - b_i v) = D^{n-1}v_i - D^{n-1}(b_i v) = a_i v - \sum_{j=0}^{n-1} \binom{n-1}{j} d^j(b_i) D^{n-1-j}(v) = a_i v - d^{n-1}(b_i) v = a_{i,0} v.$$

Hence $D^n(v_i - b_i v) = D(a_{i,0} v) = 0$ so $\{v, v_1 - b_1 v, \dots, v_{n-1} - b_{n-1} v\}$ is a set of \mathcal{K} -linearly independent vectors in $\ker(D^n)$.

Note the functor G sends V to the \mathbb{C} -vector space $W := \ker(D^n) = \text{span}_{\mathbb{C}}\{v, v_1 - b_1 v, \dots, v_{n-1} - b_{n-1} v\}$. Moreover, D induces a \mathbb{C} -linear operator N on W . By construction, this operator is nilpotent and for this basis, the matrix of N is constant (i.e., its entries belong to \mathbb{C}). Applying the functor F to (W, N) now recovers the differential module (V, D) . \square

3.5. Proof of Theorem 4 (Jordan Decomposition). The uniqueness part of the theorem is relatively easy. Since we don’t have anything new to add to Levelt’s original proof, we refer the reader to [Lev75] for the details. It remains to prove existence.

Let $D : V \rightarrow V$ be a formal differential operator. By Theorem 22, there exists a positive integer b such that $D_b : V_b \rightarrow V_b$ admits a generalised eigenspace decomposition. Thus, D_b can be represented by a block diagonal matrix where each block is upper triangular with a unique (up to similarity) eigenvalue. Thus, we may assume without the loss of generality that D_b has a unique, up to similarity, eigenvalue a . Replacing D_b by $D_b - a$, we may assume that D_b is unipotent in which case the result follows from Lemma 26. This proves the existence of Jordan decomposition for D_b .

We now show that the Jordan decomposition of D_b descends to a decomposition of D . The proof is similar to the linear setting. Picking a \mathcal{K} -basis of V and extending it to a basis of V_b allows us to write $D_b = d + A$ where A is a matrix with entries in \mathcal{K} . Let $S_b = d + B$ and $N_b = C$ for matrices B and C with respect to this basis. Then, for any $\sigma \in \text{Gal}(\mathcal{K}_b/\mathcal{K})$, it is clear that $d + A = d + \sigma(B) + \sigma(C)$ is a second Jordan decomposition of D_b . Thus, we must have $C = \sigma(C)$ and $\sigma(B) = B$. Hence, $d + B$ and C are defined over \mathcal{K} . \square

REFERENCES

- [BY15] P. Boalch and D. Yamakawa, *Twisted wild character varieties*, arXiv:1512.08091 (2015).
- [BV83] B. Babbitt and P. Varadarajan, *formal reduction theory of formal differential equations: a group theoretic view*, Pacific J. Math. **109** (1983), no. 1, 1–80.
- [Huk41] M. Hukuhara, *Théorèmes fondamentaux de la théorie des équations différentielles ordinaires. II*, Mem. Fac. Sci. Kyūsyū Imp. Univ. A. **2** (1941), 1–25.
- [KS16] M. Kamgarpour and D. Sage, *A geometric analogue of a conjecture of Gross and Reeder*, arXiv:1606.00943 (2016).
- [Kac90] V. Kac, *Infinite-dimensional Lie algebras*, 3rd ed., Cambridge University Press, Cambridge, 1990.
- [Kat70] Nicholas M. Katz, *Nilpotent connections and the monodromy theorem: Applications of a result of Turrittin*, Inst. Hautes Études Sci. Publ. Math. **39** (1970), 175–232.
- [Kat87] N. M. Katz, *On the calculation of some differential Galois groups*, Invent. Math. **87** (1987), no. 1, 13–61.
- [Ked10] K. S. Kedlaya, *p-adic differential equations*, Cambridge Studies in Advanced Mathematics, vol. 125, Cambridge University Press, Cambridge, 2010.
- [Lev75] G. Levelt, *Jordan decomposition for a class of singular differential operators*, Ark. Mat. **13** (1975), 1–27.
- [Luu15] M. Luu, *Local Langlands duality and a duality of conformal field theories*, arXiv:1506.00663 (2015).
- [Mal74] B. Malgrange, *Sur les points singuliers des équations différentielles*, Enseignement Math. (2) **20** (1974), 147–176.
- [Mal79] ———, *Sur la réduction formelle des équations différentielles à singularités irrégulières* (1979).
- [Ore33] O. Ore, *Theory of non-commutative polynomials*, Ann. of Math. (2) **34** (1933), no. 3, 480–508.
- [Pra83] C. Praagman, *The formal classification of linear difference operators*, Indagationes Mathematicae (Proceedings) **86** (1983), no. 2, 249 – 261.
- [Rob80] P. Robba, *Lemmes de Hensel pour les opérateurs différentiels. Application à la réduction formelle des équations différentielles*, Enseign. Math. (2) **26** (1980), no. 3-4, 279–311 (1981).
- [Sab] C. Sabbah, *Introduction to algebraic theory of linear systems of differential equations*, Unpublished lecture notes.
- [vdPS03] M. van der Put and M. F. Singer, *Galois theory of linear differential equations*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 328, Springer-Verlag, Berlin, 2003.
- [Tur55] H. L. Turrittin, *Convergent solutions of ordinary linear homogeneous differential equations in the neighborhood of an irregular singular point*, Acta Math. **93** (1955), 27–66.

- [Var96] V. S. Varadarjan, *Linear formal differential equations: a modern point of view*, Bulletin of AMS **33** (1996), no. 1.
- [Was65] W. Wasow, *Asymptotic expansions for ordinary differential equations*, Pure and Applied Mathematics, Vol. XIV, Interscience Publishers John Wiley & Sons, Inc., New York-London-Sydney, 1965.

E-mail address: masoud@uq.edu.au

E-mail address: samuel.weatherhog@gmail.com

SCHOOL OF MATHEMATICS AND PHYSICS, THE UNIVERSITY OF QUEENSLAND, ST. LUCIA, QUEENSLAND, AUSTRALIA